UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper 12-16-(10)
Statistics and Econometrics Series
June 2012

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

BAYESIAN MODELLING OF BACTERIAL GROWTH FOR MULTIPLE POPULATIONS

Ana P. Palacios*, J. Miguel Marín*, Emiliano Quinto and Michael P. Wiper***

Abstract:

Bacterial growth models are commonly used for the prediction of microbial safety and the shelf life of perishable foods. Growth is affected by several environmental factors such as temperature, acidity level and salt concentration. In this study, we develop two models to describe bacterial growth for multiple populations under both equal and different environmental conditions. Firstly, a semi-parametric model based on the Gompertz equation is proposed. Assuming that the parameters of the Gompertz equation may vary in relation to the running conditions under which the experiment is performed, we use feed forward neural networks to model the influence of these environmental factors on the growth parameters. Secondly, we propose a more general model which does not assume any underlying parametric form for the growth function. Thus, we consider a neural network as a primary growth model which includes the influencing environmental factors as inputs to the network. One of the main disadvantages of neural networks models is that they are often very difficult to tune which complicates fitting procedures. Here, we show that a simple, Bayesian approach to fitting these models can be implemented via the software package WinBugs. Our approach is illustrated using real experimental Listeria Monocytogenes growth data.

Keywords: Bacterial population modeling, growth functions, neural networks, Bayesian inference

* Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: appalaci@est-econ.uc3m.es, jmmarin@est-econ.uc3m.es, mwiper@est-econ.uc3m.es.
** Facultad de Medicina, Universidad de Valladolid, Palacio de Santa Cruz, Plaza de Sta. Cruz 8, 47002 Valladolid, e-mail: equinto@ped.uva.es

Bayesian modelling of bacterial growth for multiple populations

Ana P. Palacios*, Juan M. Marín*, E. Quinto** & Michael P. Wiper*

*Departamento de Estadística, Universidad Carlos III de Madrid

**Facultad de Medicina, Universidad de Valladolid

June, 2012

Abstract

Bacterial growth models are commonly used for the prediction of microbial safety and the shelf life of perishable foods. Growth is affected by several environmental factors such as temperature, acidity level and salt concentration. In this study, we develop two models to describe bacterial growth for multiple populations under both equal and different environmental conditions. Firstly, a semi-parametric model based on the Gompertz equation is proposed. Assuming that the parameters of the Gompertz equation may vary in relation to the running conditions under which the experiment is performed, we use feed forward neural networks to model the influence of these environmental factors on the growth parameters. Secondly, we propose a more general model which does not assume any underlying parametric form for the growth function. Thus, we consider a neural network as a primary growth model which includes the influencing environmental factors as inputs to the network. One of the main disadvantages of neural networks models is that they are often very difficult to tune which complicates fitting procedures. Here, we show that

a simple, Bayesian approach to fitting these models can be implemented via the software package `WinBugs`. Our approach is illustrated using real experimental *Listeria Monocytogenes* growth data.

Key words: Bacterial population modeling, growth functions, neural networks, Bayesian inference.

1 Introduction

The predictability of bacterial growth is of major interest due to the influence of bacteria on food safety and health. The evolution of microorganisms in food products can spoil the products or even cause pathogenic effects. Therefore, it is important to develop models of bacterial growth which can prevent diseases by determining the shelf life of perishable foods or by predicting the behavior of foodborne pathogens. Starting from Gompertz (1825), various so called primary, parametric, growth models which describe the evolution of the population size directly as a function of time have been developed, see e.g. McKellar and Lu (2004) for a good comparison. These models perform well in describing the evolution of bacterial density under fixed experimental conditions. Nevertheless, it is well known that bacterial growth is strongly affected by environmental conditions such as temperature, acidity or salinity of the environment and therefore, when multiple bacterial populations are analyzed, it is important to account for these effects in growth curve modeling.

In predictive microbiology, models that describe the effect of environmental conditions on the growth parameters are called secondary models, see e.g. Ross and Dalgaard (2004). For example, the square-root model of Ratkowsky et al. (1982) was developed to describe the effect of suboptimal temperature on growth rates of microorganisms. This initial approach has later been extended to include other factors such as level of acidity, water activity and salt concentra-

tion in additive or multiplicative models, see e.g. McMeekin et al. (1987), Miles et al. (1997), Wijtzes et al. (1995) and Wijtzes et al. (2001). The most common secondary models are polynomial models, see e.g. McClure et al. (1993), which allow any of the environmental factors and their interactions to be taken into account but include many parameters without biological interpretation. Another important model class is the cardinal parameter models, see Rosso et al. (1995), Augustin and Carlier (2000) and Pouillot et al. (2003) which assume that the effect of environmental factors is multiplicative.

A disadvantage of these models is that they assume simple parametric forms for the effects of the different environmental factors. Therefore, more recently, there has been interest in modeling bacteria growth curves using non-parametric approaches such as artificial neural networks, see e.g. Hajmeer et al. (1997), Geeraerd et al. (1998) and García-Gimeno et al. (2002). The advantages of neural networks are their capability to describe very complex non-linear relationships and that they do not impose any structure on the relationship between the interacting effects.

In most empirical work the fitting of any secondary models is carried out in two steps. Firstly, a primary growth model is fitted in order to estimate the growth parameters and secondly, a secondary model is fitted conditional on the estimated parameters in order to estimate the controlling factors. One problem with this strategy is that the estimated uncertainty of the first stage is not taken into account in the second stage and therefore, a poor fit at the first stage could produce inaccurate estimations at the second stage. Secondly, most work in fitting such models has used classical statistical techniques such as least squares, which, as noted in Pouillot et al. (2003), may also underestimate uncertainty. Furthermore, classical approaches do not allow for the inclusion of prior information, which can be naturally incorporated within a Bayesian

framework, see e.g. Powell et al. (2006).

In this paper, we shall develop two approaches which are applicable to growth curve estimation for bacterial populations under different environmental conditions. The first model is based on the Gompertz function where the dependence of the growth parameters on the environmental factors is modeled by a neural network. Secondly, we shall consider a direct non-parametric approach based on the use of neural networks as a primary growth model. An important feature of our approaches is that in cases where we observe bacterial growth in various colonies under possible different environmental conditions, we use hierarchical modeling to improve estimation of any single growth curve by incorporating information from the various different bacterial populations. Inference for our models is undertaken throughout using a Bayesian approach. Up to now, one problem with inference for neural networks models was that typically, complicated inference algorithms need to be designed and a great deal of tuning often needs to be carried out for these to work efficiently, see e.g. Lee (2004). Here, we show that inference can be carried out via the use of the well known `WinBugs` software through the `R2WinBugs` interface.

The rest of this paper is organized as follows. Firstly, in Section 2, we provide a brief introduction to neural networks. Then, in Section 3 we propose two alternative models for bacterial growth curves that include environmental conditions as influencing factors modeled by neural networks. In Section 4 we show how to undertake Bayesian inference for these models and then, in Section 5, we illustrate the models with an application to *listeria monocytogenes* growth curves. Finally, in Section 6, we present our conclusions and some possible extensions of our approach.

2 Feed forward neural networks

In many situations, it is assumed that there are q dependent variables, $(Y_1, \ldots, Y_q) = \mathbf{Y}$, and they can be modeled as an approximate linear or polynomial function of a set of explanatory variables, $(x_1, \ldots, x_p) = \mathbf{x}$, via e.g. multivariate regression. However, such a relationship may not always be appropriate and a more general functional relation between the dependent and independent variables must be assumed, say

$$E[\mathbf{Y}|\mathbf{x}] = \mathbf{g}(\mathbf{x})$$

where the functional form, $(g_1, \ldots, g_q) = \mathbf{g} : \mathbb{R}^p \to \mathbb{R}^q$, is unknown. One of the most popular methods of modeling the function g is via neural networks, see e.g. Stern (1996). In particular, a feed forward neural network takes a set of inputs $\mathbf{x}$, and from them computes the vector of output values as follows

$$\mathbf{g}(\mathbf{x}) = B \cdot \mathbf{\Psi}^T(\mathbf{x}^T\Gamma) \tag{1}$$

where B is a $q \times M$ matrix with $q \in \mathbb{N}$ the number of output variables and $M \in \mathbb{N}$ the number of nodes and Γ is a $p \times M$ matrix with $p \in \mathbb{N}$ being the number of explicative variables. The element $\gamma_{rk} \in \mathbb{R}$ is the weight of the connection from input r to hidden unit k and the element $\beta_{sk} \in \mathbb{R}$ is the weight connection from hidden unit k to output unit s. Finally, $\mathbf{\Psi}(a_1, \ldots, a_M) = (\Psi(a_1), \ldots, \Psi(a_M))$ where Ψ is a sigmoidal function such as the logistic function

$$\Psi(x) = \frac{exp(x)}{1 + exp(x)}, \tag{2}$$

which we will use here. Equations (1) and (2) define a feed forward neural network with logistic activation function, p explanatory variables (inputs), one hidden layer with M nodes and q dependent variables (outputs) that can be

illustrated as in Figure 1.

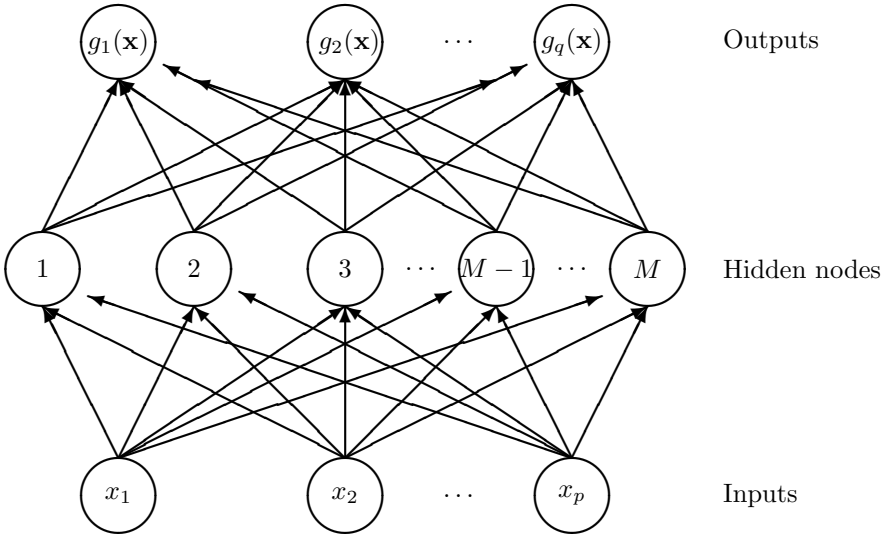

Figure 1: Neural network representation

Note that each output combines the node values in a different way. For practical fitting of neural networks models, it is typically assumed that the input variables are all defined to have a similar, finite range, e.g. $[0, 1]$. From now on, we shall assume this throughout.

3 Neural network based growth curve models

Here we develop growth curve models based on the use of neural networks to explain the functional relationship of growth to given environmental factors.

3.1 A neural network based Gompertz model

The bacterial growth process is typically characterized by three distinct phases, that is: the lag stage that reflects the adaptation of cells inoculated in a new

medium; the exponential stage that represents the bacterial growth by binary fission; and finally the stationary stage which describes the decay of the growth rate as a consequence of nutrient depletion and accumulation of waste which is followed by death or decline of the population. Sigmoidal functions which account for these three phases have been typically used to model microbial growth, see e.g. Skinner et al. (1994). In particular, the Gompertz equation is a well known model for bacterial growth over time and it has been used extensively by researchers to fit a wide variety of growth curves from different microorganisms, see e.g. Ross and McMeekin (1994) and McKellar and Lu (2004).

Here we consider a reparameterized Gompertz equation proposed by Zwietering et al. (1990). Let N_t represent the population concentration of bacteria cultivated in a Petri dish experiment at time $t \geq 0$. Then the Gompertz equation is

$$\begin{aligned} E[N_t|N_0, D, \mu, \lambda] &= g(t, N_0, D, \mu, \lambda) \quad \text{where} \\ g(t, N_0, D, \mu, \lambda) &= N_0 + D \exp\left(-\exp\left(1 + \frac{\mu e(\lambda - t)}{D}\right)\right), \end{aligned} \tag{3}$$

where e is Euler's number, N_0 is the initial bacterial density, D is the difference between the maximum bacterial density, μ is the maximum growth rate and λ is the time lag.

The primary growth model described in (3) does not allow for the case where we wish to study bacterial populations under a variety of controlled environmental conditions. Thus, suppose that we observe the growth of I bacterial populations under similar initial conditions and that we have J different environments determined by temperature, level of acidity (pH) and salt concentration (NaCI). Under fixed environmental conditions, it may be reasonable to assume that all replications have the same growth curve parameters. However, growth rates will vary under different conditions and therefore, assuming a Gompertz model,

we propose the use of neural networks to reflect the parameter dependence on the environmental factors. If N_{tij} is the concentration in population i under environmental conditions j at time t the Gompertz function is

$$E[N_{tij}|N_{0j}, D_j, \mu_j, \lambda_j] = g(t_{ij}, N_{0j}, D_j, \mu_j, \lambda_j), \tag{4}$$

where $g(\cdot)$ is as in (3), for $i = 1, \ldots, I$ and $j = 1, \ldots, J$. Now, we model the growth parameters μ, λ and D as a function of the temperature, the level of acidity and the salt concentration by a feed forward neural network, that is

$$\boldsymbol{\theta}_s = \sum_{k=1}^{M} \beta_{sk} \cdot \Psi(\mathbf{x}'\boldsymbol{\gamma}_k), \quad \text{for } s = 1, 2, 3. \tag{5}$$

where $\boldsymbol{\theta}_s$ stands for the parameters D, μ, λ and $\mathbf{x} = (T, pH, NaCI)$ is the vector of explicative variables and Ψ is the logistic function. The model defined in this section by Expression (4) and (5) will be referred to as GNN model.

3.2 A hierarchical neural network model

Here, we generalize the previous model to a new one which does not assume any underlying parametric growth function. Instead, we propose a neural network as a primary model. The output of the network is the instantaneous reproduction rate per member of the population and the inputs are the current population

size and the experimental conditions. Formally, we can write the model as

$$E[N_{tij}|N_{(t-1)ij}, f_j, T_j, pH_j, NaCl_j] =$$
$$N_{(t-1)ij} + N_{(t-1)ij} f_j(N_{(t-1)ij}, T_j, pH_j, NaCI_j) \tag{6}$$

$$f_j(N_{(t-1)ij}, T_j, pH_j, NaCI_j) =$$
$$\sum_{k=1}^{M} \beta_{jk}(\Psi(\gamma_{1k} N_{(t-1)ij} + \gamma_{2k} T_j + \gamma_{3k} pH_j + \gamma_{4k} NaCI_j)$$
$$-\Psi(\gamma_{2k} T_j + \gamma_{3k} pH_j + \gamma_{4k} NaCI_j)), \tag{7}$$

for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, $f_j(\cdot)$ is the growth rate for populations with environmental condition j. The model defined in this section by (6) will be referred to as the NN model.

3.3 Error modeling

In the previous subsections, two approaches to modeling the expected population density have been provided. These models are completed by including an error term. Thus, in the case of the full neural network model, we assume that

$$N_{tij} = N_{(t-1)ij} + N_{(t-1)ij} f_j(N_{(t-1)ij}, T_j, pH_j, NaCI_j) + \epsilon_{tij} \tag{8}$$

where we assume that the error term is

$$\epsilon_{tij}|N_{(t-1)ij}, \sigma, p \sim \mathcal{N}\left(0, \sigma^2 N_{t-1}^p\right) \tag{9}$$

where $\sigma^2 \geq 0$ and $p = 0.5$ so that the possibility that the error variance increases with population density is allowed for. Note in particular that for $p > 0$, this error structure implies that if $N_{(t-1)ij} = 0$, then $N_{tij} = 0$, so that once the

population has died out, then it remains extinct. Figure 2 illustrates different bacterial growth curves from petri dish experiments under the same conditions. It can be seen that the curves are closer together initially when the population density is lower and diverge over time as the population density grows which suggests that a model of this type is reasonable. Following the same idea of increasing error variance we assume for the GNN model that the error term is

$$\epsilon_{tij}|gt_{ij}, \sigma, p \sim \mathcal{N}\left(0, \sigma^2 g(t_{ij})^p\right) \tag{10}$$

where $g(.)$ is the Gompertz function evaluated at the current time point.

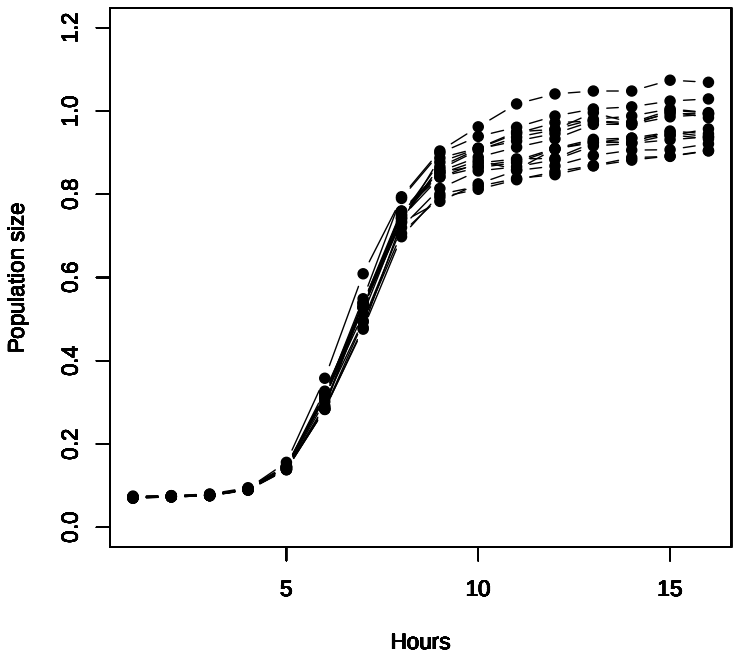

Figure 2: 15 replications of bacterial growth under $T = 42°\text{C}$, $pH = 7.4$ and $NaCI = 2.5\%$.

4 Bayesian inference for the neural network models

Given a set of observed inputs and outputs from a neural network, say $D = (x_1, y_1), \ldots, (x_N, y_N)$, inference can be carried out using a variety of approaches, see e.g. Neal (1996) and Fine (1999) for reviews. Here, we shall consider a Bayesian approach. In order to implement such an approach, we must first define suitable prior distributions for the neural network parameters β and γ and for the uncertainty. Firstly, we suppose little prior knowledge concerning the variance and hence we propose a vague, inverse-gamma, prior distribution for it $\sigma^{-2} \sim G(a/2, b/2)$. In neural network models is common to use relative uninformative prior distributions due to the scarcity of prior information about the parameters. For simplicity we choose normal and gamma distributions with hierarchical structure, that is

$$\begin{aligned} \beta_{ik}|m_{i\beta}, \sigma_\beta^2 &\sim \mathcal{N}\left(m_{i\beta}, \sigma_\beta^2\right) \\ \boldsymbol{\gamma}_k|m_\gamma, \sigma_\gamma^2 &\sim \mathcal{N}\left(\boldsymbol{m}_\gamma, \sigma_\gamma^2 I\right), \end{aligned}$$

where the subscript i in the GNN model accounts for the growth parameters and in the NN model for the groups defined by the environmental conditions. The Bayesian approach is completed by vague, but proper prior distributions

for the remaining hyperparameters as follows:

$$\begin{aligned}
m_{i\beta}|\sigma_\beta^2 &\sim \mathcal{N}\left(m_{0\beta}, \frac{\sigma_\beta^2}{c_\beta}\right) \\
m_{0\beta}|\sigma_\beta^2 &\sim \mathcal{N}\left(0, \frac{\sigma_\beta^2}{e_\beta}\right) \\
\frac{1}{\sigma_\beta^2} &\sim \mathcal{G}\left(\frac{d_{\beta 1}}{2}, \frac{d_{\beta 2}}{2}\right) \\
\boldsymbol{m}_\gamma|\sigma_\gamma^2 &\sim \mathcal{N}\left(\mathbf{0}, \frac{\sigma_\gamma^2}{c_\gamma} I\right) \\
\frac{1}{\sigma_\gamma^2} &\sim \mathcal{G}\left(\frac{d_{\gamma 1}}{2}, \frac{d_{\gamma 2}}{2}\right),
\end{aligned}$$

where c_β, e_β, $d_{\beta 1}$, $d_{\beta 2}$, c_γ, $d_{\gamma 1}$ and $d_{\gamma 2}$ are assume known and fixed. Similar hierarchical prior distributions are typically used in Bayesian inference for neural network models, see e.g. Lavine and West (1992), Müller and Insua (1998) and Andrieu et al. (2001). For alternatives, see e.g. Lee (2004), Robert and Mengersen (1999) and Roeder and Wasserman (1997).

Usually, we will have good prior knowledge about the average initial population density, $m_0 = E[N0_i|m_0, s_0]$ and the variance, s_0, as typically, petri dishes are seeded with very similar quantities of bacteria close to a known, theoretical level, so we shall typically assume that these are known. Otherwise, a simple non-informative prior distribution $f(m_0, t_0) \propto 1/t_0$, where $t_0 = 1/s_0^2$ can be used when, immediately, we have that given the observed set of initial densities, $\mathbf{N0} = (N0_1, \ldots, N0_I)$,

$$\begin{aligned}
m_0|\mathbf{N0}, s_0 &\sim \mathcal{N}\left(\overline{N0}, \frac{s_0^2}{I}\right) \\
s_0^2|\mathbf{N0} &\sim \mathcal{IG}\left(I-1, \sum_{i=1}^{I}(N0_i - \overline{N0})^2\right)
\end{aligned}$$

where $\overline{N0} = \frac{1}{I}\sum_{i=1}^{I} N0_i$ is the average initial density.

Given the above prior structure, a closed form for the posterior parameter distributions is not available. However, Markov-Chain Monte-Carlo (MCMC) techniques can be employed to allow us to generate an approximate Monte Carlo sample from the posterior parameter distributions, see e.g. Gilks et al. (1996) for a full review. Various different MCMC algorithms have been proposed in the neural networks literature, but in general the efficiency of such samplers depends on the model, see e.g. Lee (2004).

As an alternative, here, we propose using the generic MCMC sampler, `WinBugs`, as developed by Spiegelhalter et al. (1999), which is appropriate for hierarchical modeling situations, programmed in combination with `R`, via `R2WinBugs`.

Figure 3 illustrates the dependence structure of the NN model in `WinBugs` style (although code cannot be constructed directly from this diagram). In the figure, random and logical nodes are represented by ellipses and fixed nodes (independent variables) are represented by rectangles. The arrows represent dependence relationships with the single arrows showing stochastic dependence and the double arrows representing logical dependence. For more details see `http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml`.

As `WinBugs` is a generic approach to MCMC sampling, it is important to check on the convergence of the sampler. Various tools can be used to check the convergence. In particular, as well as standard graphical techniques such as looking at the trace, the evolution of the mean and the autocorrelations of the sampled output, we also use formal diagnostic techniques such as the modified Gelman-Rubin statistic, as in Brooks and Gelman (1998).

4.1 Model selection

Thus far, inference is conditional on the number of hidden nodes, M, being unknown. Various approaches to estimating M may be considered. One pos-

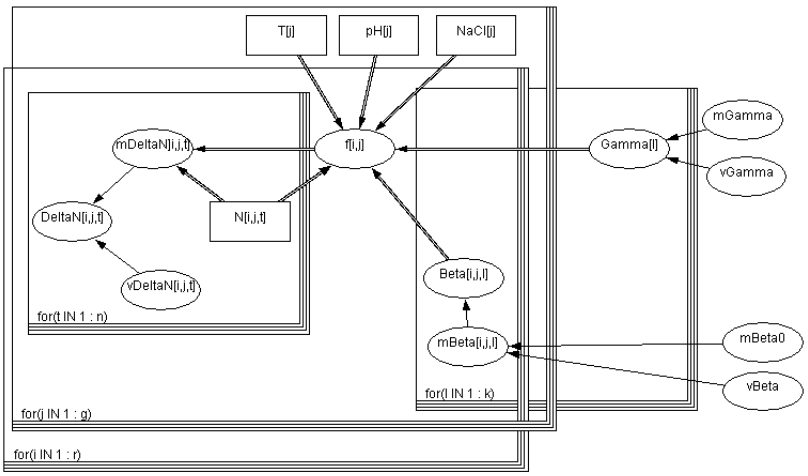

Figure 3: Dependence structure of the NN model

sibility is to treat M as a variable and given a prior distribution for M, use variable dimensional MCMC approaches to carry out inference, see e.g. Müller and Insua (1998) or Neal (1996). Another approach which we shall use in this article is to use an appropriate model selection technique to choose the value of M.

A number of criteria have been proposed for model selection in Bayesian inference. A standard, Bayesian selection criterion which is particularly appropriate when inference is carried out using MCMC methods is the deviance information criterion (DIC), as proposed in Spiegelhalter et al. (2002). However, in the context of neural networks, the possible lack of identifiability of the model or multimodality of the posterior densities make this criterium unstable. Many variants of the DIC have also been considered and here, we prefer to apply the DIC_3 criterion of Celeux et al. (2006). For a model $\mathcal{M}$ with parameters $\boldsymbol{\theta}$ and observed data $\mathbf{y}$ the DIC_3 is defined as follows:

$$DIC_3 = -4E_{\boldsymbol{\theta}}[\log f(\mathbf{y}|\boldsymbol{\theta})|\mathbf{y}] + 2\log \prod_{i=1}^{n} E_{\boldsymbol{\theta}}[f(y_i|\boldsymbol{\theta},\mathbf{y})].$$

In Celeux et al. (2006) this criterion is recommended in the context of latent variable models.

An alternative approach which we also consider when comparing different models is the posterior predictive loss performance (PPLP) proposed by Gelfand and Ghosh (1998). Based on the posterior predictive distribution, this criterion consists in defining a weight loss function which penalizes actions for departure from the corresponding observed value as well as for departure from what we expect the replication to be. In this way, the approach is a compromise between the two types of departures: fit and smoothness. For squared error loss, the criterion becomes

$$PPLP = \frac{k}{k+1}\sum_{i=1}^{n}(m_i - y_i)^2 + \sum_{i=1}^{n} s_i^2$$

, where $m_i = E[y_i^{rep}|\mathbf{y}]$ and $s_i^2 = Var[y_i^{rep}|\mathbf{y}]$ are, respectively, the mean and the variance of the predictive distribution of y_i^{rep} given the observed data $\mathbf{y}$ and k is the weight we assign to departures from the observed data. The first term of the $PPLP$ is a plain goodness-of-fit term and the second term penalizes complexity and rewards parsimony.

5 Application: Listeria monocytogenes

In this section we analyze a data set consisting of measures of the concentrations of Listeria monocytogenes bacteria in a petri dish under several experimental conditions. The environmental factors taken into account are temperature, level of acidity and salinity. Temperatures range between 22°C and 42°C, pH between 4.5 and 7.4 and NaCI between 2.5% and 5.5%. There are 96 different combinations of environmental factors (we call groups) and for each group there are several replications (between 15 and 20, depending on the group). The number

of observations per curve varies between 16 and 24, depending on the curve. We kept for the analysis 74 groups (excluding the cases with extreme values of factors which inhibit growth) and chose randomly 10 replications for each one.

Using the DIC_3 criterion as outlined earlier, the optimum number of nodes for both models is 2. Temperature, pH and NaCI as inputs of the neural networks were previously scaled onto [0.1, 0.9]. In the implementation of the GNN model we keep the hyperparameters $m_{i\beta}$, σ_β, m_γ and σ_γ fixed at $m_{i\beta} = 0$, $\sigma_\beta = 10$, $\boldsymbol{m}_\gamma = (0, \ldots, 0)'$ and $\sigma_\gamma = 10$. Regarding the error variance we choose $a = 0.2$ and $b = 0.2$. In the NN model the highest level of hyperparameters were set to $c_\beta = 10$, $e_\beta = 10$, $d_{\beta 1} = 0.1$, $d_{\beta 2} = 0.01$, $c_\gamma = 10$, $d_{\gamma 1}$ and $d_{\gamma 2} = 0.01$.

In order to fit the models, in each case we generated chains with random initial values and 200000 iterations each, including 100000 iterations of burn-in. To diminish autocorrelation between the generated values we also used a thinning rate of 1000. Trace plots and autocorrelation functions were used to check convergence in the predictions and in all cases it was found that the burn-in period of 100000 iterations was reasonable. Furthermore, the Gelman-Rubin statistic was equal o very close to 1 for predictions, being a good indicator of convergence.

In order to have a benchmark for the comparison of models we also fit two different simple models, the independent Gompertz model and the pooled Gompertz model. The first one implies that each observed curve, including the replications, is independent and therefore has its own Gompertz growth parameters. Independent, relatively diffuse normal $N(0, 100)$ prior distributions are assumed for these parameters. In contrast, the pooled model assumes that the replications under a fixed set of environmental conditions are samples from a unique, underlying growth curve for that set of conditions. Normal priors are then placed on the parameters of this growth curve as for the independent

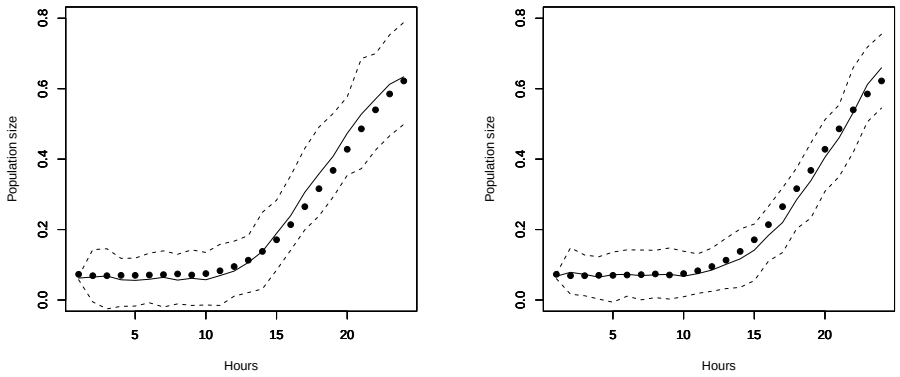

Figure 4: Fitting bacterial growth curves

model. For both benchmark models the errors are the same as in the GNN case with a $\mathcal{G}(0.1, 0.1)$ prior distribution for the error variance.

The $DIC3$ and the $PPLP$ criteria were computed in order to compare the different models under consideration and Table 1 shows the estimated values for all of these models. As is expected, the pooled model performs better than de independent one since the assumption of independence for all the curves is somewhat extreme. Therefore, it seems reasonable to assume different curves under different environmental conditions but under equal conditions we assume a common curve. And this is the approach we choose for the proposed models. But the problem of this model is that it does not explain the effect of the environmental factors and it is needed to estimate one model for each group of conditions. Then, regarding our proposed models which incorporate the environmental factors as explicative variables the results show that hierarchical neural network model outperforms the Gompertz model with neural networks for the parameters. The $DIC3$ and the $PPLP$ values are lowest for the former model.

Figure 4 shows for a particular curve ($T = 34°\text{C}$, $pH = 6.5$ and $NaCI =$

5.5%) the fitting of both models. On the left, the Gompertz model with neural networks explaining the dependence of the growth parameter on the environmental factor and on the right the fitting of the hierarchical neural network model. The observed values are represented by points, the estimated growth curves are represented by the solid line, and the dashed lines represents the 95% credible interval computed from the posterior distributions. It can be observed that the fit is good in both cases and in GNN model the credible interval included all the true observations. Note that in NN model three observations do not fall within the credible interval due to the overestimation of the lag period. In the remaining curves (replications and different group conditions), we also found good fits for both models. Similar results are observed in the fitted plots for all the groups.

Table 1: Model comparison

Model	DIC3	PPLP
Independet Gompertz	-19136	781
Pooled Gompertz	-39420	211
Gomp & NN	-40099	41
Neural Networks	-58492	28

Now, we consider one-step ahead predictions. That is, for a particular curve we observe data until observation t and predict the population size at $t+1$. In the next step, we observe data until $t+1$ and predict the population size at $t+2$ and so on, until the completion of the predictive curve. Figure 5 shows the one-step-ahead predictive curves for both models for a particular growth curve ($T = 42°$C, $pH = 5.5$ and $NaCI = 2.5\%$). In contrast with the fit, the Gompertz model shows a better predictive performance. The mean square error of the prediction in the Gompertz model is equal to 0.001, while for the NNs

model is 0.008. But in the second model higher accuracy is reached as can be seen from the narrower credible interval.

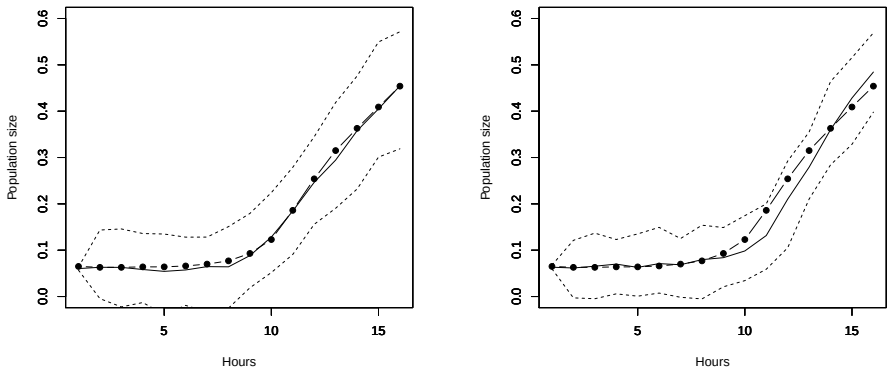

Figure 5: One-step ahead predictions

In the context of model checking, several authors, e.g. Gelfand (1996) and Vehtari and Lampinen (2003) have proposed the use of cross-validatory predictive densities. Following this approach, the data is divided in two subsets $(\mathbf{y_1}, \mathbf{y_2})$. The first of these is used to fit the model and to estimate the posterior distribution of the parameters, while the second set is used to compute the cross-validatory predictive density: $f(\mathbf{y_1}|\mathbf{y_2}) = \int f(\mathbf{y_2}|\boldsymbol{\theta}) f(\boldsymbol{\theta}|\mathbf{y_1}) d\boldsymbol{\theta}$. In our case, we computed the predictive density for one of the groups which was not used in the model fitting. The environmental conditions for this new group are $T = 26°\text{C}$, $pH = 6.5$ and $NaCI = 5.5\%$. Figure 6 shows the mean prediction (solide line) and the 95% credible interval (dashed line) for both models, GNN on the left and NN on the right. As there are many replications for this group, we plot only the mean curve and shade the area between the minimum value and the maximum value observed for each time t among replications. As an input of the neural network for the NN model we used the mean curve of the replications.

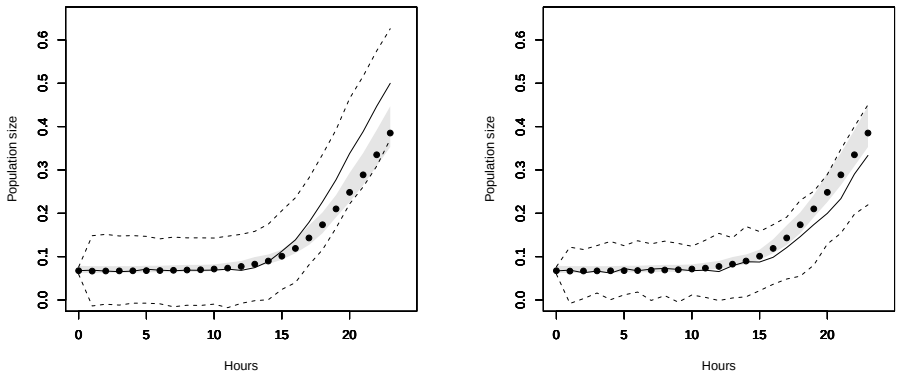

Figure 6: Cross-validation

Although both cross-validation predictions are fairly good, in the case of the GNN model some observations lie outside the credible interval. Moreover, comparing the mean prediction with the mean observed curve, the NN model yields more accurate predictions.

6 Conclusions and extensions

In this article we have illustrated that neural networks can be used to model bacterial growth for multiple populations. Neural networks were used as a secondary model that explains the dependence on environmental factors and also as a primary model which, besides time, includes experimental conditions as explicative variables. Inference was carried on in a Bayesian approach that avoids the problems for doing inference in two steps. Both models yield accurate estimations and good predictions which show that NNs can be used to model bacterial growth describing accurately the complex interacting effects of environmental factors without imposing any simplifying assumption.

Estimations were implemented in `WinBugs` via `R2WinBugs` showing that `WinBugs`

can be a powerful and flexible tool able to handle very complex models such as neural networks with great ease. As MacKay (1995) pointed out, Gibbs sampling method is not the most efficient of MCMC methods, but there may be problems of interest where the convenience of this tool outweighs this drawback.

On the other hand, the modified Gompertz equation was used as the base model for the first approach we considered but other parametric bacteria growth models such as Baranyi or logistic are equally applicable.

A restriction in the models as assumed here is that we suppose that data are equally spaced in time. Although this is typically the case in petri dish experiments, this may not be true with more general populations. In the case of irregularly spaced data, differential equation models with diffusion type approximations with the neural network models for the growth functions may be considered (see Donnet et al. (2010)).

Finally, alternative approximations to the neural network models for growth functions may be considered as spline methods from a classical point of view or the use of gaussian process approximations.

Acknowledgements

Support for this work was provided by the Spanish Ministry of Education and Science.

References

Andrieu, C., N. Freitas, and A. Doucet (2001). Robust full Bayesian learning for radial basis networks. *Neural Computation 13*(10), 2359–2407.

Augustin, J. and V. Carlier (2000). Modelling the growth rate of Listeria monocytogenes with a multiplicative type model including interactions between

environmental factors. *International journal of food microbiology 56*(1), 53–70.

Brooks, S. and A. Gelman (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics 7*(4), 434–455.

Celeux, G., F. Forbes, C. Robert, and D. Titterington (2006). Deviance information criteria for missing data models. *Bayesian Analysis 1*(4), 651–674.

Donnet, S., J. Foulley, and A. Samson (2010). Bayesian analysis of growth curves using mixed models defined by stochastic differential equations. *Biometrics 66*(3), 733–741.

Fine, T. (1999). *Feedforward neural network methodology*. Springer Verlag.

García-Gimeno, R., C. Hervás-Martínez, et al. (2002). Improving artificial neural networks with a pruning methodology and genetic algorithms for their application in microbial growth prediction in food. *International journal of food microbiology 72*(1-2), 19.

Geeraerd, A., C. Herremans, C. Cenens, and J. Van Impe (1998). Application of artificial neural networks as a non-linear modular modeling technique to describe bacterial growth in chilled food products. *International journal of food microbiology 44*(1-2), 49–68.

Gelfand, A. (1996). Model determination using sampling-based methods. *WGilks, SRichardson, DSpiegelhalter (eds.), Markov Chain Monte Carlo in Practice*, 145–161.

Gelfand, A. and S. Ghosh (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika 85*(1), 1.

Gilks, W., S. Richardson, and D. J. Spiegelhalter (1996). *Markov chain Monte Carlo in practice*. London: Chapman and Hall.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London 115*, 513–583.

Hajmeer, M., I. Basheer, and Y. Najjar (1997). Computational neural networks for predictive microbiology II. Application to microbial growth. *International journal of food microbiology 34* (1), 51–66.

Lavine, M. and M. West (1992). A Bayesian method for classification and discrimination. *Canadian Journal of Statistics 20* (4), 451–461.

Lee, H. (2004). *Bayesian nonparametrics via neural networks*. Society for Industrial Mathematics.

MacKay, D. (1995). Bayesian methods for neural networks: Theory and applications. Technical report, Cambridge: Cavendish Laboratory, Cambridge University.

McClure, P., J. Baranyi, E. Boogard, T. Kelly, and T. Roberts (1993). A predictive model for the combined effect of pH, sodium chloride and storage temperature on the growth of Brochothrix thermosphacta. *International journal of food microbiology 19* (3), 161–178.

McKellar, R. and X. Lu (2004). *Modeling microbial responses in food*. CRC Press.

McMeekin, T., R. Chandler, P. Doe, C. Garland, J. Olley, S. Putro, and D. Ratkowsky (1987). Model for combined effect of temperature and salt

concentration/water activity on the growth rate of Staphylococcus xylosus. *Journal of Applied Microbiology 62*(6), 543–550.

Miles, D., T. Ross, J. Olley, and T. McMeekin (1997). Development and evaluation of a predictive model for the effect of temperature and water activity on the growth rate of Vibrio parahaemolyticus. *International journal of food microbiology 38*(2-3), 133–142.

Müller, P. and D. R. Insua (1998). Issues in Bayesian analysis of neural network models. *Neural Computation 10*(3), 749–770.

Neal, R. (1996). *Bayesian learning for neural networks*. Springer Verlag.

Pouillot, R., I. Albert, M. Cornu, and J. Denis (2003). Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to Listeria monocytogenes. *International journal of food microbiology 81*(2), 87–104.

Powell, M., M. Tamplin, B. Marks, and D. Campos (2006). Bayesian synthesis of a pathogen growth model: Listeria monocytogenes under competition. *International journal of food microbiology 109*(1-2), 34–46.

Ratkowsky, D., J. Olley, T. McMeekin, and A. Ball (1982). Relationship between temperature and growth rate of bacterial cultures. *Journal of Bacteriology 149*(1), 1.

Robert, C. and K. Mengersen (1999). Reparameterisation issues in mixture modelling and their bearing on MCMC algorithms. *Computational Statistics and Data Analysis 29*(3), 325–344.

Roeder, K. and L. Wasserman (1997). Practical Bayesian density estimation using mixtures of normals. *Journal of the American Statistical Association 92*(439), 894–902.

Ross, T. and P. Dalgaard (2004). *Secondary models*. In Modeling microbial responses in food. CRC Press.

Ross, T. and T. McMeekin (1994). Predictive microbiology. *International journal of food microbiology 23*(3-4), 241–264.

Rosso, L., J. Lobry, S. Bajard, and J. Flandrois (1995). Convenient model to describe the combined effects of temperature and pH on microbial growth. *Applied and Environmental Microbiology 61*(2), 610.

Skinner, G., J. Larkin, and E. Rhodehamel (1994). Mathematical modeling of microbial growth: a review. *Journal of food safety 14*(3), 175–217.

Spiegelhalter, D., N. Best, B. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B, Statistical Methodology 64*(4), 583–639.

Spiegelhalter, D., A. Thomas, and N. Best (1999). WinBUGS version 1.2 user manual. *MRC Biostatistics Unit*.

Stern, H. (1996). Neural networks in applied statistics. *Technometrics 38*(3), 205–214.

Vehtari, A. and J. Lampinen (2003). Expected utility estimation via cross-validation. *Bayesian statistics 7*, 701–710.

Wijtzes, T., J. De Wit, et al. (1995). Modelling bacterial growth of Lactobacillus curvatus as a function of acidity and temperature. *Applied and environmental microbiology 61*(7), 2533.

Wijtzes, T., F. Rombouts, M. Kant-Muermans, K. Van't Riet, and M. Zwietering (2001). Development and validation of a combined temperature, water activity, pH model for bacterial growth rate of Lactobacillus curvatus. *International journal of food microbiology 63*(1-2), 57–64.

Zwietering, M., I. Jongenburger, F. Rombouts, and K. Van't Riet (1990). Modeling of the bacterial growth curve. *Applied and Environmental Microbiology 56*(6), 1875.